Выделение движущихся объектов при наличии масштабных преобразований в последовательности изображений

П.В. Бабаян I , Е.С. Кожина I

¹ ФГБОУ ВО РГРТУ им. В.Ф. Уткина, , ул. Гагарина, 59/1, Рязань 390005, Россия

Аннотация

В работе рассматривается задача выделения подвижных объектов методом вычитания фона при наличии масштабных геометрических преобразований. В качестве алгоритма оценки геометрических преобразований масштаба используется многоэталонный алгоритм. Представлены результаты экспериментальных исследований, проведенных с использованием натурных видеосюжетов, которые в процессе тестирования подвергались влиянию масштабирования с помощью программного обеспечения. В ходе тестирования алгоритма исследовалось влияние параметра СКО оценки масштаба на качество выделения объектов.

Предложенный алгоритм показывает достаточно высокую эффективность применения по сравнению с существующими алгоритмами. Он может быть использован в системах технического зрения летательных аппаратов, мобильных роботов, в системах видеонаблюдения.

Ключевые слова

Выделение объекта, оценка геометрических преобразований, разностный алгоритм, масштабные преобразования, многоэталонный алгоритм.

Extraction of Moving Objects in the Presence of Scale Transformations in a Sequence of Images

P.V. Babayan ¹, E.S. Kozhina ¹

Abstract

The paper considers the problem of extraction of moving objects by subtracting the background in the presence of scale geometric transformations. A multi-reference algorithm is used as an algorithm for estimation of scale transformations. An algorithm for automatic extraction of moving objects is proposed, taking into account the geometric transformations caused by the movement of the image sensor.

The results of experimental studies obtained using real-world videos affected by scaling using software. During the testing of the algorithm, the influence of the scale estimation parameter on the quality of object extraction was investigated.

The algorithm proposed shows sufficiently higher efficiency in comparison with commonly used approach. The algorithm can be used in machine vision systems of aircrafts, mobile robots, and video surveillance systems.

Keywords

Object extraction, estimation of geometric transformations, subtraction algorithm, scale transformations, multi-template algorithm.

1. Введение

В настоящее время актуальной задачей обработки и анализа последовательностей изображений является выделение движущихся объектов. Разработано огромное количество бортовых систем,

ORCID: 0000-0003-4041-3101 (П.В. Бабаян¹); 0000-0003-4498-2694 (Е.С. Кожина²);

© 2022 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

19–22 September 2022, Ryazan, Russia

¹ Ryazan State Radio Engineering University named after V.F. Utkin, Gagarina str., 59/1, Ryazan 390005, Russia

ГрафиКон 2022: 32-я Международная конференция по компьютерной графике и машинному зрению, 19-22 сентября 2022 г., Рязанский государственный радиотехнический университет им. В.Ф. Уткина, Рязань, Россия EMAIL: pvbradio@yandex.ru (П.В. Бабаян¹); kozhina.e.s@rsreu.ru (Е.С. Кожина¹);

решающих данную задачу, которые предназначены для установки на автономных носителях: самолётах, вертолётах, БПЛА, автомобилях и др. Важными характеристиками таких систем является необходимость работы в реальном масштабе времени, и эти системы должны работать самостоятельно, требуя лишь незначительных усилий оператора. За многие годы разработано достаточно большое количество алгоритмов, которые позволяют решать такие задачи, но большинство из них работает качественно только при определенных условиях. К основным факторам, которые затрудняют выделение объектов, можно отнести следующее: движения датчика в пространстве, такие как поворот, смещение и масштабные искажения, наличие турбулентности атмосферы, отсутствие информации об объекте. Соответственно, применение существующих алгоритмов в таких условиях иногда бывает практически невозможным без дополнительных эвристик.

Именно поэтому важной остается проблема разработки систем выделения объектов в видеопоследовательностях, сформированных при наличии геометрических преобразований. Исходя из этого, сформулируем цель исследования, которая состоит в разработке методов выделения движущихся объектов в сложных условиях наблюдения, характеризующихся геометрическими трансформациями масштаба и изменениями фонового изображения.

Задача алгоритма выделения движения состоит в принятии решения о наличии или отсутствии объекта для каждой точки изображения. Результат работы такого алгоритма в конечном итоге представляет собой бинарную маску, состоящую из единичных и нулевых точек, где единице соответствуют точки, относящиеся к объекту, а нулевые точки относятся к фону.

В результате работы алгоритма выделения движения получается бинарная маска, которая состоит из единичных и нулевых точек. Единичным значениям соответствуют точки, где находится объект, а нулевым – фон.

Идеальным результатом служит полное выделение объекта и отсутствие ложно отмеченных единичных точек на всем изображении. Как правило, бинарная маска все же содержит ошибочно выделенные точки, которые отнесены к объекту и не содержит части точек, на самом деле принадлежащих объекту [1].

Существующие методы, решающие задачу выделения объектов, можно разделить на следующие группы [2]:

- 1. Статистические методы;
- 2. Методы на основе машинного обучения;
- 3. Методы на основе пространственно-временной фильтрации.

Рассмотрим первую группу методов. Работа **статистических методов** выделения объектов основана на использовании априорной информации о различии статистических свойств объекта и фона. Среди методов, которые относятся к данной группе, можно выделить подходы на базе пороговой обработки, наращивания областей, байесовской сегментации.

Если говорить о преимуществах данных методов, то это, безусловно, выигрыш в сложности вычисления операций, а также возможность работы алгоритма при перемещающемся датчике изображений. Среди недостатков таких алгоритмов отмечается главное условие — а именно необходимость статистического различия точек фона и объекта. К сожалению, этот недостаток не даёт возможности использовать их для решения поставленной задачи, потому что в данной работе исходными данными могут быть видеопоследовательности со сложными условиями наблюдения [3].

Методы на основе машинного обучения относятся к методам искусственного интеллекта, которые позволяют построить обучаемые модели. В настоящее время эти подходы активно развиваются и все больше находят свое практическое применение, но главным условием работы данных методов, является то, что перед тем, как запустить процесс машинного обучения, необходимо загрузить в компьютер значительное количество исходных данных, на которых алгоритм будет учиться обрабатывать запросы. После процесса обучения программа уже сама сможет распознавать объекты на новых изображениях [4].

Методы выделения объектов на основе **пространственно-временной** фильтрации комбинируют в себе 2 отдельно существующих метода пространственной фильтрации и временные методы.

Пространственные методы основаны на использовании различных операций линейной и нелинейной пространственной фильтрации наблюдаемых изображений. Для выделения малоразмерных объектов используется медианная и мультимедианная фильтрация, для выделения границ объектов - градиентные маски.

Недостатком методов пространственной фильтрации является то, что для их использования необходимо различие в свойствах пространственной структуры объекта и фона.

Временные методы [5] используют информацию, получаемую в предыдущих кадрах. Эффективность таких методов, безусловно, выше, чем у методов, использующих при обнаружении только априорную информацию о пространственных и статистических свойствах фона и объекта. С точки зрения потребностей в априорной информации, данная методика лучше всего соответствует задаче, поставленной в настоящей работе.

Таким образом, разумным решением является объединение преимуществ двух последних методов. В результате объединения получился пространственно-временной метод, который исключает недостатки объединяемых методов.

Во многих работах предложен и развит пространственно-временной подход к решению задачи выделения и оценки параметров движущихся объектов. Особенностью является использование модели перемещения объекта на изображении. Алгоритм содержит, помимо этапов уточнения яркостей фонового изображения и вычисления апостериорных вероятностей, этап корреляционного поиска эталонного изображения на текущем изображении.

Существующие методы решения задачи выделения объектов в основном ориентированы на работу при наблюдении сцены с помощью неподвижного датчика и в отсутствии геометрических изменений изображения. В случае, когда получаемые изображения испытывают геометрические трансформации, разработанные алгоритмы выдают не совсем желаемый результат. Из-за присутствующих искажений очень сложно а, бывает, совершенно невозможно обнаружить движущиеся объекты. Поэтому в данной работе авторы предлагают алгоритм, который учитывает наличие геометрических преобразований.

Так как исходными данными является только последовательность изображений без какой-либо другой информации, целесообразно использовать пространственно-временные методы, но так как данные методы работают хорошо только при неподвижном фоне, нужно сначала достаточно оценить параметры геометрических преобразований и по этим оценкам стабилизировать фон.

От того, насколько точно оцениваются геометрические преобразования, будет зависеть результат выделения объектов. Полученные оценки используются для компенсации геометрических преобразований фона на наблюдаемых изображениях. После компенсации можно переходить к выделению объектов.

Среди алгоритмов оценки параметров геометрических преобразований к рассматриваемой задаче в наибольшей степени подходят следующие:

- 1. Многоэталонный алгоритм [2];
- 2. Метод на основе преобразований Фурье [6].

В основе многоэталонного алгоритма лежит принцип, построенный на выборе нескольких опорных участков на изображении и определении их положения от кадра к кадру. Для слежения за опорными участками используется алгоритм оценки положения на основе сопоставления с эталоном [2, 5].

Чтобы оценить параметры геометрических преобразований по методу преобразований Фурье [6], нужно использовать теоремы о смещении, вращении и масштабировании для спектров функций двух переменных. Таким образом, при сдвиге, повороте и масштабировании функции происходит такой же поворот и масштабирование амплитудного спектра, а сдвиг изображения не оказывает влияния на амплитудный спектр функции.

Важным моментом при разработке алгоритма является простота реализации в бортовых вычислительных комплексах, именно поэтому в настоящей работе для оценки параметра масштаба используется многоэталонный алгоритм, так как алгоритмы на основе преобразований Фурье обладают сложной программно-аппаратной реализацией.

2. Алгоритм выделения объектов

В настоящей работе для выделения объектов предлагается использовать пространственновременной алгоритм, который относится к классу методов на основе вычитания фона [1, 7].

В условиях геометрических преобразований этот метод необходимо дополнить процедурой оценки и компенсации геометрических преобразований изображения.

Для начала определимся с математической моделью формирования изображения и введем критерий качества выделения, а затем выясним, какое влияние оказывает точность оценки параметров геометрических преобразований на качество выделения объектов. В качестве меры точности оценки параметров примем величину дисперсии ошибки оценивания.

Модель формирования наблюдаемого изображения без учёта эффектов дискретизации выглядит следующим образом:

$$l(x, y) = g(\Omega(x, y))(1 - r(x, y)) + h(x, y)r(x, y) + \xi(x, y),$$
(1)

где l(x,y) – наблюдаемое изображение, g(x,y) – фоновое изображение, $\Omega(x,y)$ – двумерная функция, описывающая геометрические преобразования, которые испытывает фоновое изображение, r(x,y) – бинарное изображение, единичные значения которого задают расположение объектов, h(x,y) – яркостное изображение всех объектов, $\xi(x,y)$ – аддитивный шум датчика.

В самом простейшем случае изменения масштаба можно описать выражением:

$$S(x,y) = (\lambda x, \lambda y), \tag{2}$$

где λ – коэффициент изменения масштаба.

Если g(x,y) в окрестности точки (\hat{x},\hat{y}) меняется медленно, то одним из подходов к упрощению процедуры выделения может служить линейная интерполяция фонового изображения в окрестности точки (\hat{x},\hat{y}) . Применяя разложение Тейлора в окрестности точки (\hat{x},\hat{y}) , имеем

$$g(\mu_{x},\mu_{y}) \approx g(\hat{x},\hat{y}) + \frac{\partial g(\hat{x},\hat{y})}{\partial x} (\hat{\lambda}y'\Delta\varphi - \hat{\lambda}\zeta_{x} + x'\Delta\lambda - \Delta\alpha) + \frac{\partial g(\hat{x},\hat{y})}{\partial y} (-\hat{\lambda}x'\Delta\varphi - \hat{\lambda}\zeta_{y} + y'\Delta\lambda - \Delta\beta),$$
(3)

где $\mu_x = \hat{x} + \hat{\lambda} y' \Delta \varphi - \hat{\lambda} \zeta_x + x' \Delta \lambda - \Delta \alpha$ и $\mu_y = \hat{y} - \hat{\lambda} x' \Delta \varphi - \hat{\lambda} \zeta_y + y' \Delta \lambda - \Delta \beta$, $\hat{\lambda}$, $\hat{\varphi}$, $\hat{\alpha}$, $\hat{\beta}$ – оценки измнения параметров масштаба, поворота и смещения по x и y, соответственно, $\hat{x} = \hat{\lambda} x' - \hat{\alpha}$ и $\hat{y} = \hat{\lambda} y' - \hat{\beta}$, $x' = x_0 \cos \hat{\varphi} + y_0 \sin \hat{\varphi}$ и $y' = -x_0 \sin \hat{\varphi} + y_0 \cos \hat{\varphi}$, x_0 , y_0 – координаты точки фонового изображения, ζ_x , ζ_y – случайные величины, которые описывают геометрические деформации, вызванные атмосферной турбулентностью. Величины $\Delta \alpha$, $\Delta \beta$, $\Delta \varphi$, $\Delta \lambda$ – это отклонения между истинными значениями параметров преобразований и их оценками. Они являются нормальными, центрированными и независимыми случайными величинами с СКО, равными $\sigma_{\alpha\beta}$, $\sigma_{\alpha\beta}$, σ_{φ} , σ_{z} , соответственно.

Приводя подобные слагаемые, получим

$$g(\mu_{x},\mu_{y}) \approx g(\hat{x},\hat{y}) + \hat{\lambda} \left(\frac{\partial g(\hat{x},\hat{y})}{\partial x} y' - \frac{\partial g(\hat{x},\hat{y})}{\partial y} x' \right) \Delta \varphi - \frac{\partial g(\hat{x},\hat{y})}{\partial x} (\hat{\lambda}\zeta_{x} + \Delta \alpha) - \frac{\partial g(\hat{x},\hat{y})}{\partial y} (\hat{\lambda}\zeta_{y} + \Delta \beta) + \left(\frac{\partial g(\hat{x},\hat{y})}{\partial x} x' + \frac{\partial g(\hat{x},\hat{y})}{\partial y} y' \right) \Delta \lambda.$$

$$(4)$$

Вычислим математическое ожидание и дисперсию для $f = g(\mu_x, \mu_y)$, как для линейной комбинации нормально распределённых случайных величин:

$$E(f) = g(\hat{x}, \hat{y}) \tag{5}$$

$$D(f) = \hat{\lambda}^2 \left[\nabla g(\hat{x}, \hat{y}) \begin{pmatrix} y' \\ -x' \end{pmatrix} \right]^2 \sigma_{\varphi}^2 + \left| \nabla g(\hat{x}, \hat{y}) \right|^2 (\hat{\lambda}^2 \sigma_z^2 + \sigma_{\alpha\beta}^2) + \left[\nabla g(\hat{x}, \hat{y}) \begin{pmatrix} y' \\ x' \end{pmatrix} \right]^2 \sigma_{\lambda}^2, \tag{6}$$

где $\nabla g(\hat{x},\hat{y}) = \left(\frac{\partial g}{\partial x},\frac{\partial g}{\partial y}\right)$ – градиент яркости изображений, $\hat{\sigma}_{_{\varphi}}^{_{2}}$ – дисперсия оценки угла поворота, $\sigma_{_{z}}^{^{2}}$

– оценка СКО геометрических деформаций, $\sigma_{\alpha\beta}^{2}$ – дисперсии оценки смещения по x и y, полагаемых равными одной и той же величине, σ_{α}^{2} – дисперсия оценки масштаба.

Учитывая, что распределение p(l/r=0) является нормальным с известной дисперсией, нахождение порогового значения Λ' эквивалентно заданию $(1-p_{_-})\cdot 100\%$ доверительного интервала для распределения p(l/r=0).

Для принятия решения о наличии объекта на изображении используется критерий Неймана-Пирсона. В соответствии с данным критерием принимается решение о наличии объекта, если величина

l не попадает в $(1-p_{-})\cdot 100\%$ доверительный интервал для распределения p(l/r=0). В противном случае принимается решение о присутствии фона.

Таким образом, зная яркость фонового изображения g(x,y), получая оценки $(\hat{\alpha},\hat{\beta},\hat{\phi},\hat{\lambda})$, дисперсии σ_z^2 , $\sigma_{\alpha\beta}^2$, σ_φ^2 , σ_ζ^2 , σ_ζ^2 , можно произвести оптимальное по критерию Неймана–Пирсона выделение объекта путём проверки попадания наблюдаемого значения яркости l в $(1-p_{\perp})\cdot 100\%$ доверительный интервал яркости, центрированный на величине $g(\hat{x},\hat{y})$.

В присутствии случайных геометрических деформаций задание доверительного интервала производится с учётом градиента наблюдаемого изображения. На качественном уровне это можно объяснить следующим образом. При деформациях изображения наибольшие ошибки измерения яркости будут присутствовать вблизи резких перепадов яркостей. Поэтому в этих областях ширину доверительного интервала надо увеличить. Напротив, в тех местах, где изображение однородно, деформации не вызывают больших ошибок, и даже небольшое изменение яркости свидетельствует о наличии объекта.

Формально правило принятия рещения можно записать так:

$$\hat{r} = \begin{cases} 1, \text{ если } |g(\hat{x}, \hat{y}) - l| > d(p_{-}) \sqrt{D(l/r = 0)}; \\ 0, \text{ иначе,} \end{cases}$$
 (7)

где $d(p_{-})$ — полуширина $(1-p_{-})\cdot 100\%$ доверительного интервала для нормированной гауссовой случайной величины, дисперсию D(l/(r=0)) можно найти из выражения

$$D(l/(r=0) = \hat{\lambda}^{2} \left[\nabla g(\hat{x}, \hat{y}) \begin{pmatrix} y' \\ -x' \end{pmatrix} \right]^{2} \sigma_{\varphi}^{2} + \left| \nabla g(\hat{x}, \hat{y}) \right|^{2} (\hat{\lambda}^{2} \sigma_{z}^{2} + \sigma_{\alpha\beta}^{2}) + \left[\nabla g(\hat{x}, \hat{y}) \begin{pmatrix} y' \\ x' \end{pmatrix} \right]^{2} \sigma_{\lambda}^{2} + \sigma_{\xi}^{2}, \quad (8)$$

где $\hat{\lambda}^2$ — оценка изменения масштаба, $\nabla g(\hat{x},\hat{y})$ — градиент яркости изображения, $\hat{\sigma}_{\xi}^2$ — оценка дисперсии аддитивного шума. Фактически (x',y') является вектором расстояния между точкой (\hat{x},\hat{y}) и центром вращения фонового изображения.

Если смещение и поворот равны нулю, то формула (8) примет вид

$$D(l/(r=0) = \left|\nabla g(\hat{x}, \hat{y})\right|^2 (\hat{\lambda}^2 \sigma_z^2) + \left[\nabla g(\hat{x}, \hat{y}) \begin{pmatrix} y' \\ x' \end{pmatrix}\right]^2 \sigma_{\lambda}^2 + \sigma_{\xi}^2. \tag{9}$$

Дальнейшее выделение объектов производится с применением выражений (7) и (9).

3. Экспериментальные исследования

В ходе экспериментальных исследований исследовалось влияние параметра СКО оценки масштаба σ_{λ} на качество выделения объектов.

Моделирование работы алгоритма проводилось на натурных видеосюжетах, снятых на камеру видимого диапазона. Преобразования масштаба имитировались с помощью программного обеспечения.

На рисунке 1 приведены кадры из исходной видеопоследовательности (a), соответствующее ему эталонное бинарное изображение (б) и результаты выделения объекта предложенным алгоритмом при значении параметра $\sigma_{\lambda}^2 = 0$ (в); $\sigma_{\lambda}^2 = 0,0001$ (г); $\sigma_{\lambda}^2 = 0,0002$ (д); $\sigma_{\lambda}^2 = 0,0003$ (е).

Идеальным результатом работы алгоритма является выделение всех точек объекта и безошибочное определение фона (рисунок 1(б)), но в реальных условиях, разумеется, такого варианта не бывает. На бинарном изображении, кроме объекта, будут присутствовать точки фона, которые были ложно отнесены к объекту (рисунок 1 (в-е)).

На рисунке 1 (в) приведен кадр одного из сюжетов при значении параметра $\sigma_{\lambda}^2 = 0$. По сути это результат работы уже существующего алгоритма без учета масштаба. Результат предложенного алгоритма, который учитывает изменение масштаба на видео, представлен на рисунке 1 под буквами г-е. Визуально можно сделать вывод, что учет в предложенном алгоритме параметра СКО изменения масштаба положительно влияет на качество выделения объектов, ложно выделенных точек стало намного меньше. Стоит также отметить, что значение σ_{λ}^2 подбиралось путем экспериментов. По

представленным изображениям видно, что при σ_{λ}^2 =0,0001 и σ_{λ}^2 =0,0002 наблюдается повышение качества работы алгоритма, дальнейшее увеличение этого значения влечет за собой ухудшение качества выделения.

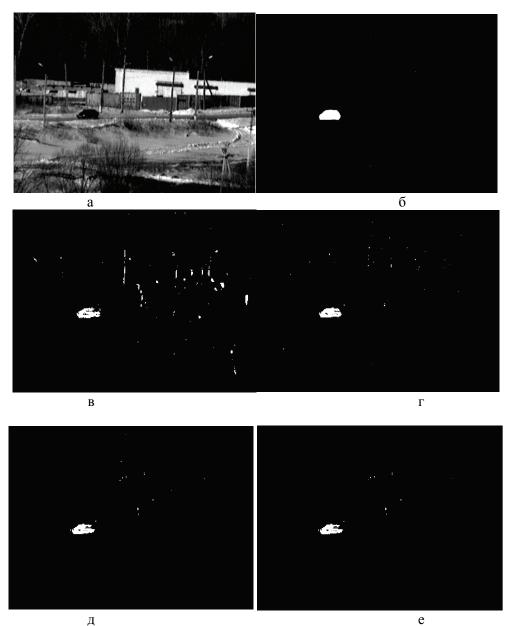


Рисунок 1 – Кадр исходной видеопоследовательности (а), соответствующее ему эталонное бинарное изображение (б) и результаты выделения объектов предложенным алгоритмом при значении параметра $\sigma_{_{\lambda}}{}^{2}$ =0 (в), $\sigma_{_{\lambda}}{}^{2}$ =0,0001 (г), $\sigma_{_{\lambda}}{}^{2}$ =0,0002 (д), $\sigma_{_{\lambda}}{}^{2}$ =0,0003 (е)

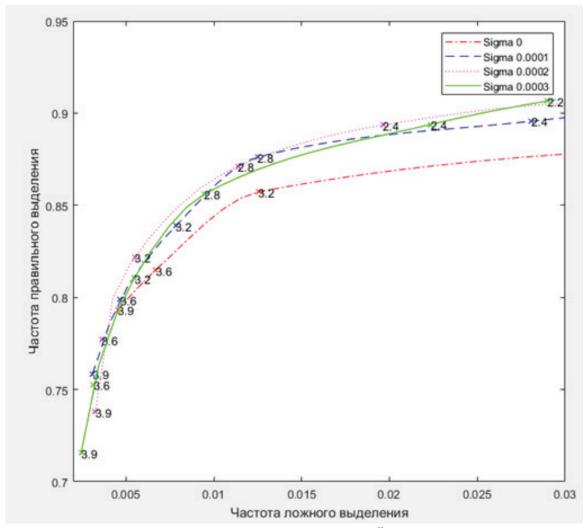


Рисунок 1 - ROC-кривая для экспериментальных исследований

Из графика, представленного на рисунке 2, видно, что при учете параметра СКО масштаба ($\sigma_{_{\lambda}}^{^{2}}$) график смещается вверх и влево. Это означает, что качество выделения объектов выше, чем было раньше. Ложно выделенных точек стало меньше (значение FP меньше у предложенного алгоритма), а количество правильно выделенных точек при этом возросло (значение FP выше у предложенного алгоритма).

Также по графику можно отметить, что при значении частоты правильного выделения 0,86, качество выделения объектов при учете параметра масштаба σ_{λ}^{2} =0,0002 выросло в 1,5 раза (на 35%), чем при σ_{λ}^{2} =0.

4. Заключение

В данной работе предложен новый подход к выделению объектов, который учитывает влияние масштабирования изображения в алгоритме на основе вычитания фона. Благодаря тому, что в формуле оценок дисперсий учитывается параметр СКО изменения масштаба, происходит более качественное выделение объектов.

Проведены сравнительные экспериментальные исследования предложенного алгоритма, а также определены частоты правильного и ложного выделения объектов. Экспериментальные исследования проводились на нескольких различных видеосюжетах, отличающихся между собой фоновой обстановкой и количеством объектов в кадре. Экспериментальные исследования разработанного алгоритма показали эффективность его применения и подтвердили, что учет параметра СКО ошибки оценивания масштаба, положительно влияет на качество работы алгоритма. Удалось повысить качество работы алгоритма в 1,5 раза (на 35%). Заметно уменьшилось количество ложных

срабатываний. В дальнейшем планируется провести эксперименты на видеосюжетах, подверженных одновременно влиянию смещения, масштабирования и поворота датчика изображения в пространстве, а также расширить выборку сюжетов, используемых при экспериментальных исследованиях.

5. Список источников

- [1] Бабаян П.В., Кожина Е.С. Автоматическое выделение объектов в видеоинформационной системе в условиях масштабных преобразований // Цифровая обработка сигналов и её применение. DSPA-2022: Доклады 24-ой Международной конференции, Москва: РНТОРЭС им. А.С. Попова, 2022. С. 257-260.
- [2] Hakim Bendjenna, Abdallah Meraoumia, Othaila Chergui, «Pattern recognition system: from classical methods to deep learning techniques», J. Electron. Imaging 27(3), 033008 (2018).
- [3] Junhua Yan, Jun Du, Yong Young, «Real-time unmanned aerial vehicle tracking of fast moving small target on ground», J. Electron. Imaging 27(5), 053010 (2018).
- [4] Xiaohe Wu, Weisong Wang, Fei Yang, «Joint learning of Siamese network with top-down modulation and hard example mining for visual tracking», J. Electron. Imaging 28(5), 053034 (2019).
- [5] Danelljan, A. Robinson, F. S. Khan, and M. Felsberg, «Beyond correlation filters: Learning continuous convolution operators for visual tracking», in Proc. Eur. Conf. Comput. Vis., 2016, pp. 472–488.
- [6] Дискретное преобразование Фурье в обработке сигналов: учеб. пособие / Н. В. Воропаева, С. Я. Новиков, М. Е. Федина. Самара: Изд-во «Самарский университет», 2015. 48 с.
- [7] Zhang Y. J. Handbook of Image Engineering // Справочник по имиджевой инженерии. Springer, Singapore, 2021. C. 1963.